JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 18, 355-364 (1967)

Duality for Variational Problems

BERTRAM MOND

Aerospace Research Laboratories,

Wright-Patterson Air Force Base, Ohio 45433

AND

MORGAN A. HANSON

Queens University, Kingston, Ontario, and

Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio 45433

Submitted by Richard Bellman

INTRODUCTION

In [1], Courant and Hilbert, quoting an earlier work of Friedrichs [2], give a dual relationship for a simple type of unconstrained variational problem. More recently Hanson [3], Kreindler [4], Pearson [5] and [6], and Ringlee [7] have given dual formulations for different forms of the control problem.

In this note, we give a dual formulation for a class of variational problems with differential inequality constraints. Both fixed end point and free boundary value problems are considered. Our results include a converse duality theorem-theorem 3, which is the main result of this paper.

Hanson [3] pointed out that some of the duality theorems of mathematical programming have analogues in the variational calculus. This relationship between mathematical programming and the classical calculus of variations is explored and extended. In essence, it is shown that our duality theorems can be considered as dynamic generalizations of corresponding (static) duality theorems of mathematical programming. Since mathematical programming and the classical calculus of variations have undergone independent development, it is felt that the mutual adaption of ideas and techniques may prove fruitful.

DUALITY

Let $f(t, x, x')$ be a (real) scalar function with continuous derivatives up to and including the second order with respect to each of its arguments and let

$Q(t, x, x')$ be an m-dimensional function which similarly has continuous derivatives up to and including the second order. x is an n-dimensional function of t; prime denotes derivative with respect to t.

Consider the determination of a piecewise smooth extremal $x = x(t)$, $t_0 \leqslant t \leqslant t_1$, for the following modified Lagrange problem:

Problem I (Primal) $\equiv P$

$$\text{Minimize} \int_{t_0}^{t_1} f(t, x, x')\, dt \tag{1}$$

$$\text{Subject to } x(t_0) = x_0, \qquad x(t_1) = x_1 \tag{2}$$

and

$$Q(t, x, x') \geqslant 0. \tag{3}$$

Consider also the determination of an $n + m$ dimensional extremal $(x, \lambda) = (x(t), \lambda(t))$, $t_0 \leqslant t \leqslant t_1$, for the following maximization problem:

Problem II (Dual) $\equiv D$

$$\text{Maximize} \int_{t_0}^{t_1} \{f(t, x, x') - \lambda(t)\, Q(t, x, x')\}\, dt \tag{4}$$

$$\text{Subject to } x(t_0) = x_0, \qquad x(t_1) = x_1 \tag{5}$$

$$f_x(t, x, x') - \lambda(t)\, Q_x(t, x, x') = \frac{d}{dt}\, [f_{x'}(t, x, x') - \lambda(t)\, Q_{x'}(t, x, x')] \tag{6}$$

$$\lambda(t) \geqslant 0. \tag{7}$$

Here $x(t)$ is an n-dimensional piecewise smooth function and $\lambda(t)$ is an m-dimensional function continuous except possibly for values of t corresponding to corners of $x(t)$. For values of t corresponding to corners of $x(t)$, (6) must be satisfied for unique right- and left-hand limits.

No notational distinction is made between row and column vectors. Subscripts denote partial derivatives; superscripts denote vector components. Thus

$$f_x - \lambda Q_x = \frac{d}{dt}\, [f_{x'} - \lambda Q_{x'}]$$

means

$$f_{x^1} - \Big(\sum_{i=1}^{m} \lambda^i Q^i\Big)_{x^1} = \frac{d}{dt}\Big[f_{x^{1\prime}} - \Big(\sum_{i=1}^{m} \lambda^i Q^i\Big)_{x^{1\prime}}\Big]$$

$$f_{x^2} - \Big(\sum_{i=1}^{m} \lambda^i Q^i\Big)_{x^2} = \frac{d}{dt}\Big[f_{x^{2\prime}} - \Big(\sum_{i=1}^{m} \lambda^i Q^i\Big)_{x^{2\prime}}\Big]$$

$$\vdots \qquad\qquad \vdots$$

$$f_{x^n} - \Big(\sum_{i=1}^{m} \lambda^i Q^i\Big)_{x^n} = \frac{d}{dt}\Big[f_{x^{n\prime}} - \Big(\sum_{i=1}^{m} \lambda^i Q^i\Big)_{x^{n\prime}}\Big].$$

THEOREM 1. *If f is convex and Q concave in x and x', then the infimum of P is greater than or equal to the supremum of D.*

PROOF. Let $(x^*, x^{*\prime})$ satisfy (2) and (3) and let (x, x', λ) satisfy (5), (6), and (7). Then

$$\int_{t_0}^{t_1} \{f(t, x^*, x^{*\prime}) - f(t, x, x')\}\, dt$$

$$\geqslant \int_{t_0}^{t_1} \{(x^* - x) f_x(t, x, x') + (x^{*\prime} - x') f_{x'}(t, x, x')\}\, dt \text{ (since } f \text{ is convex)}$$

$$= \int_{t_0}^{t_1} (x^* - x) \left\{ \frac{d}{dt} [f_{x'}(t, x, x') - \lambda(t) Q_{x'}(t, x, x')] + \lambda(t) Q_x(t, x, x') \right\} dt$$

$$+ \int_{t_0}^{t_1} (x^{*\prime} - x') f_{x'}(t, x, x')\, dt \qquad \text{(by (6))}$$

$$= (x^* - x) \{f_{x'}(t, x, x') - \lambda(t) Q_{x'}(t, x, x')\} \Big|_{t=t_0}^{t=t_1}$$

$$+ \int_{t_0}^{t_1} \lambda(t) \{(x^* - x) Q_x(t, x, x') + (x^{*\prime} - x') Q_{x'}(t, x, x')\}\, dt$$

(by integration by parts)

$$= \int_{t_0}^{t_1} \lambda(t) \{(x^* - x) Q_x(t, x, x') + (x^{*\prime} - x') Q_{x'}(t, x, x')\}\, dt \quad \text{(by (2) and (5))}$$

$$\geqslant \int_{t_0}^{t_1} \lambda(t) \{Q(t, x^*, x^{*\prime}) - Q(t, x, x')\}\, dt \qquad \text{(since } Q \text{ is concave)}$$

$$\geqslant \int_{t_0}^{t_1} - \lambda(t) Q(t, x, x')\, dt \qquad \text{(by (7) and (3)).}$$

Hence

$$\int_{t_0}^{t_1} f(t, x^*, x^{*\prime})\, dt \geqslant \int_{t_0}^{t_1} \{f(t, x, x') - \lambda(t) Q(t, x, x')\}\, dt,$$

and, therefore, the infimum of P is greater than or equal to the supremum of D.

The *convexity* of f and the *concavity* of Q with respect to x and x' will henceforth be assumed.

The following necessary conditions for the existence of an extremal of P are given by Valentine [8]. We assume throughout that necessary constraints for the existence of multipliers are satisfied.

For every minimizing arc $x = x^*(t)$ of P, there exists a function of the form

$$F = \lambda_0^* f - \lambda^*(t) Q \tag{8}$$

such that

$$F_x = \frac{d}{dt}(F_{x'}) \tag{9}$$

$$\lambda^{*i} Q^i = 0, \qquad i = 1,..., m \tag{10}$$

$$\lambda^*(t) \geqslant 0 \tag{11}$$

hold throughout $t_0 \leqslant t \leqslant t_1$ (except at corners of $x^*(t)$ where (9) holds for unique left and right hand limits). Here λ_0^* is a constant, $\lambda^*(t)$ is continuous except possibly for values of t corresponding to corners of $x^*(t)$, and $(\lambda_0^*, \lambda^*(t))$ cannot vanish for any t, $t_0 \leqslant t \leqslant t_1$.

It will be assumed that the minimizing arc $x^*(t)$ is *normal*, i.e., that λ_0^* can be taken equal to 1.

THEOREM 2. *If the function $x^*(t)$ minimizes the primal problem P, then there exists a $\lambda^*(t)$ such that $(x^*(t), \lambda^*(t))$ maximizes the dual problem D and the extreme values of P and D are equal.*

PROOF. Since $x^*(t)$ minimizes P, it follows from the results of Valentine that there exists a $\lambda^*(t)$ such that (6) and (7) hold. Thus $(x^*(t), \lambda^*(t))$ satisfies the constraints of D.

In addition, we have, from (10)

$$\lambda^*(t) Q(t, x^*, x^{*\prime}) = 0. \tag{12}$$

(12) and Theorem 1 imply that $(x^*(t), \lambda^*(t))$ is a maximizing solution of D.

CONVERSE DUALITY

We now consider the converse dual problem, that is, of finding conditions under which the existence of a maximizing solution to problem D implies the existence of a minimizing solution to problem P. The situation is complicated by the fact that the constraints of D are of a different form from the constraints in the Valentine problem.

The constraint (6) can be written

$$\begin{aligned} G(t, x, x', x'', \lambda, \lambda') \equiv (f_x - \lambda Q_x) - (f_{x't} - \lambda Q_{x't}) + \lambda' Q_{x'} \\ - (f_{x'} - \lambda Q_{x'})_x x' - (f_{x'} - \lambda Q_{x'})_{x'} x'' = 0. \end{aligned} \tag{13}$$

We assume that G is twice continuously differentiable with respect to each of its arguments.

Second order derivatives with respect to t can be eliminated from (13) by the introduction of additional variables. Thus, letting $z = x'$, where $z \equiv (z_1, z_2, \dots, z_n)$ is an n-dimensional vector function of t, (13) becomes

$$G(t, x, x', z', \lambda, \lambda') = 0 \tag{14}$$
$$x' = z.$$

The theorems of Valentine [8] giving necessary conditions for the existence of an extremal can now be applied to D. z can then be eliminated from the resulting necessary conditions by the reintroduction of x''. When this is done, the following necessary conditions are obtained:

For every maximizing arc $(x, \lambda) \equiv (x^*(t), \lambda^*(t))$ of D, where $(x^{*\prime}(t), \lambda^*(t))$, is piecewise smooth, there exists a function

$$H \equiv \mu_0^*(f - \lambda^* Q) - \mu^*(t)\, G - \nu^*(t)\, \lambda^* \tag{15}$$

such that the conditions

$$H_x - \frac{d}{dt} H_{x'} + \frac{d^2}{dt^2} H_{x''} = 0 \tag{16}$$

$$H_\lambda = \frac{d}{dt} H_{\lambda'} \tag{17}$$

$$\nu^{*i}\lambda^{*i} = 0, \qquad i = 1, \dots, m \tag{18}$$

$$\nu^*(t) \leqslant 0 \tag{19}$$

hold except at corners of $(x^{*\prime}(t), \lambda^*(t))$ where (16) and (17) hold for unique right and left hand limits. Here μ^* and ν^* are, respectively, n and m dimensional functions of t, continuous except possibly for values of t corresponding to corners of $(x^{*\prime}(t), \lambda^*(t))$. μ_0^* is a constant; μ_0^*, $\mu^*(t)$, and $\nu^*(t)$ cannot vanish simultaneously for any t, $t_0 \leqslant t \leqslant t_1$.

(16) is the usual generalization of the Euler-Lagrange differential equations when higher order derivatives appear (see, e.g., [1], p. 190). The reversal of the inequality in (19) is due to the fact that we are now dealing with a maximization rather than a minimization problem.

Note that D is a problem with free boundary values for x' (or z) and for λ. As is well-known (see, e.g., [1] 208-211) the necessary conditions for the existence of an extremal remain valid.

We shall assume, from now on, that the maximizing solution of D is *normal*, i.e. that μ_0^* may be taken equal to 1.

THEOREM 3. *If* (x^*, λ^*) *is a maximizing function for* D *such that* $(x^{*\prime}, \lambda^*)$ *is piecewise smooth in* $t_0 \leqslant t \leqslant t_1$, *and such that*

$$\mu(t)\, G_x - \frac{d}{dt}[\mu(t)\, G_{x'}] + \frac{d^2}{dt^2}[\mu(t)\, G_{x''}] = 0 \tag{20}$$

only has the solution $\mu(t) = 0$, $t_0 \leqslant t \leqslant t_1$, *then* $x^*(t)$ *is a minimizing solution of* P, *and the extreme values of* P *and* D *are equal.*

PROOF. It follows from the results of Valentine that there exists a $\mu^*(t)$ and $\nu^*(t)$ such that along the arc (x^*, λ^*), $t_0 \leqslant t \leqslant t_1$,

$$f_x - \lambda^* Q_x - \mu^*(t)\, G_x - \frac{d}{dt}\left[f_{x'} - \lambda^* Q_{x'} - \mu^*(t)\, G_{x'}\right] + \frac{d^2}{dt^2}\left[-\mu^*(t)\, G_{x''}\right] = 0 \tag{21}$$

$$-Q - \mu^*(t)\, G_\lambda - \nu^* = \frac{d}{dt}\left[-\mu^*(t)\, G_{\lambda'}\right] \tag{22}$$

$$\nu^* \leqslant 0 \tag{23}$$

$$\nu^{*i} \lambda^{*i} = 0, \qquad i = 1,\ldots, m. \tag{24}$$

(6) and (21) imply (20). Since $\mu^*(t) = 0$, $t_0 \leqslant t \leqslant t_1$, (22) becomes

$$-Q - \nu^* = 0. \tag{25}$$

(23) and (25) imply

$$Q \geqslant 0.$$

Thus $x^*(t)$ satisfies the constraints of P. From (24) and (25) it follows that

$$\lambda^* Q = 0. \tag{26}$$

It follows from (26) and Theorem 1 that $x^*(t)$ is a minimizing solution of P. (20) is of the form

$$A(t)\frac{d^2\mu}{dt^2} + B(t)\frac{d\mu}{dt} + C(t)\,\mu = 0, \tag{27}$$

where A, B, and C are $n \times n$ matrix functions of t. In general, a solution of a system of second-order differential equations contains two arbitrary constant vectors. Since μ is not required to fulfill any boundary conditions, the requirement that $\mu = 0$ be the only solution cannot, in general, be fulfilled. If, however, $A(t)$ and $B(t)$ are simultaneously zero for all t, $t_0 \leqslant t \leqslant t_1$, then (27) reduces to

$$C(t)\,\mu(t) = 0.$$

(This will be the case, for example, if f and Q are both linear in x'.) $\mu = 0$ is then the only solution of (27) if $C(t)$ is nonsingular for all t, $t_0 \leqslant t \leqslant t_1$.

Natural Boundary Values

It is possible to extend the duality theorems established in the previous two sections to the corresponding variational problem with natural boundary values rather than fixed end points.

(Primal) $\equiv P$

$$\text{Minimize} \int_{t_0}^{t_1} f(t, x, x')\, dt$$

$$\text{Subject to } Q(t, x, x') \geqslant 0.$$

(Dual) $\equiv D$

$$\text{Maximize} \int_{t_0}^{t_1} \{f(t, x, x') - \lambda(t) Q(t, x, x')\}\, dt$$

$$\text{Subject to}$$

$$f_x(t, x, x') - \lambda(t) Q_x(t, x, x') = \frac{d}{dt}\,[f_{x'}(t, x, x') - \lambda(t) Q_{x'}(t, x, x')]$$

$$\lambda(t) \geqslant 0$$

$$[f_{x'}(t, x, x') - \lambda(t) Q_{x'}(t, x, x')]_{t=t_0} = 0 \tag{28}$$

and

$$[f_{x'}(t, x, x') - \lambda(t) Q_{x'}(t, x, x')]_{t=t_1} = 0. \tag{29}$$

By $[f_{x'} - \lambda Q_{x'}]_{t=t_i}$ we mean the n-dimensional vector $f_{x'} - \lambda Q_{x'}$ evaluated at $t = t_i$.

We shall not repeat the proofs of Theorems 1-3, but will merely point out the modifications in the arguments that are required for the theorems to remain valid.

In the proof of Theorem 1, (2) and (5) were utilized to assure that the term $(x^* - x)(f_{x'} - \lambda Q_{x'})\,|_{t=t_0}^{t=t_1}$ is zero. It is obvious that even though (2) and (5) are lacking in the problems with natural boundary values, the term $(x^* - x)(f_{x'} - \lambda Q_{x'})\,|_{t=t_0}^{t=t_1}$ still vanishes by virtue of (28) and (29).

It is well known (see, e.g., [1] 208-211) that for a problem with natural boundary values one has, in addition to the necessary conditions for the corresponding fixed end point problem, the natural boundary conditions $F_{x'}\,|_{t=t_0} = 0$ and $F_{x'}\,|_{t=t_1} = 0$, where F is the appropriate Lagrangian function. In our problem, this means that with F defined as in (8), one has as necessary conditions, (28) and (29) in addition to (9), (10), and (11). Thus, as before, if x^* is a minimizing solution of P, there exists a λ^* such that (x^*, λ^*) satisfies the constraints of D; Theorem 2 then follows from Theorem 1 and (10).

Finally, with (28) and (29) added to the constraints of D, the argument justifying Theorem 3 can be appropriately modified by adding

$$-\beta_0[f_{x'} - \lambda^* Q_{x'}]_{t=t_0} - \beta_1[f_{x'} - \lambda^* Q_{x'}]_{t=t_1},$$

β_0 and β_1 constants, to the function H in (15). Since the necessary conditions (16), (17), (18), and (19) are unchanged, Theorem 3 remains valid.

If only one end point is fixed, say $x(t_0) = x_0$, then the corresponding boundary condition (28) is omitted. The discussion in this section is easily modified to show that Theorems 1-3 are also valid in this case.

Mathematical Programming

Consider problems P and D with natural boundary values when all functions are independent of t. (For simplicity, take $t_1 - t_0 = 1$.) P and D reduce to

Problem 1* (Primal) P^*

$$\text{Minimize } f(x)$$

$$\text{Subject to } Q(x) \geqslant 0.$$

Problem 2* (Dual) D^*

$$\text{Maximize } f(x) - \lambda Q(x)$$

$$\text{Subject to } f_x(x) - \lambda Q_x(x) = 0$$

$$\lambda \geqslant 0.$$

Problems 1* and 2* are just the dual programs treated in [9], [10], and [11]. Indeed, each of our 3 duality theorems has its counterpart in the literature of mathematical programming.

The proof in [9] of the counterpart to our Theorem 2 is based on the Kuhn-Tucker conditions [12]. These conditions state that if x^* is a minimizing solution of Problem 1* there exists a λ^* such that

$$f_x(x^*) - \lambda^* Q_x(x^*) = 0 \tag{30}$$

$$\lambda^* \geqslant 0 \tag{31}$$

$$\lambda^{i*} Q^i(x^*) = 0, \qquad i = 1,\dots, m. \tag{32}$$

Thus the Kuhn-Tucker conditions (30), (31), and (32) may be regarded as the static counterparts of the corresponding Euler-Lagrange and Clebsch conditions (9), (11), and (10).

Huard [9], Mangasarian [10], and Wolfe [11] make the assumption that the constraints satisfy the Kuhn-Tucker Constraint Qualification [12]. Correspondingly, we assume normality. Cottle [13] points out that the Kuhn-Tucker Constraint Qualification "has the effect of assuring the existence of multipliers with $\lambda_0^* = 1$," i.e. of assuring normality. The relationship between the Kuhn-Tucker Constraint Qualification and conditions for normality is discussed in [14] and [15]. Hadley [16] gives necessary and sufficient conditions for x^* to be a normal solution of problem 1* and actually derives the Kuhn-Tucker conditions based on assumptions that assure that x^* is normal.

The condition for converse duality in Theorem 3, i.e., that $\mu(t) = 0$ be the only solution of (20), reduces in the static case to the requirement that

$$[f_{xx}(x^*) - \lambda^* Q_{xx}(x^*)]\,\mu = 0$$

only have the solution $\mu = 0$. This is, of course, equivalent to the requirement that the matrix $f_{xx}(x^*) - \lambda^* Q_{xx}(x^*)$ be nonsingular-which is just the condition for converse duality given in [9], [10], and [11].

Various extensions of the duality theorems of non-linear programming have appeared recently in the literature. Although we shall not pursue this matter in this paper, in many instances, corresponding extensions of our variational duality theorems are possible.

References

1. R. Courant and D. Hilbert. "Methods of Mathematical Physics," Vol. 1. Wiley (Interscience), New York, 1943.
2. K. O. Friedrichs. Ein Verfahren der Variations-Rechnung das Minimum eines Integral als das Maximum eines Anderen Ausdruckes Dazustellen. Göttingen Nachrichten, 1929.
3. M. A. Hanson. Bounds for functionally convex optimal control problems. *J. Math. Anal. Appl.* **8** (1964), 84-89.
4. E. Kreindler. Reciprocal optimal control problems. *J. Math. Anal. Appl.* **14** (1966), 141-152.
5. J. D. Pearson. Reciprocity and duality in control programming problems. *J. Math. Anal. Appl.* **10** (1965), 383-408.
6. J. D. Pearson. Duality and a decomposition technique. *SIAM J. Control.* **4** (1966), 164-172.
7. A. V. Ringlee. Bounds for convex variational programming problems arising in power system scheduling and control. *IEEE Trans. Aut. Cont.* **1965**, 28-35.

8. F. A. VALENTINE. The problem of Lagrange with differential inequalities as added side conditions. "Contributions to the Calculus of Variations 1933-37," pp. 407-448. Univ. of Chicago Press, 1937.
9. P. HUARD. Dual programs. *IBM J. Res. Dev.* 6 (1962), 137-139.
10. O. L. MANGASARIAN. Duality in nonlinear programming. *Quart. Appl. Math.* 20 (1962), 300-302:
11. P. WOLFE. A duality theorem for non-linear programming. *Quart. Appl. Math.* 19 (1961), 239-244.
12. H. W. KUHN AND A. W. TUCKER. Nonlinear programming. *Proc. 2nd Berkeley Symp. on Mathematical Statistics and Probability*, pp. 481-492. Univ. of Calif. Press, 1951.
13. R. W. COTTLE. Nonlinear programs with positively bounded Jacobians, ORC 64-12 (RR). Operations Research Center, University of California, Berkeley, Calif., 1964.
14. K. J. ARROW AND L. HURWICZ. Reduction of constrained maxima to saddle-point problems. *Proc. 3rd Berkeley Symp. on Mathematical Statistics and Probability*, Vol. V, pp. 1-20. Univ. of California Press, 1955.
15. K. J. ARROW, L. HURWICZ, AND H. UZAWA. Constraint qualifications in maximization problems. *Naval Res. Logist. Quart.* 8 (1961), 175-191.
16. G. HADLEY. "Nonlinear and Dynamic Programming." Addison-Wesley, Reading, Mass., 1964.